

Jet properties at high-multiplicity

Erik Gerwick^{1,*} and Peter Schichtel^{2,3,†}

¹*II. Physikalisches Institut, Universität Göttingen, Germany*

²*Institut für Theoretische Physik, Universität Heidelberg, Germany, and*

³*Institute for Particle Physics Phenomenology, Durham University, UK*

We investigate the behaviour of jets at high-multiplicity using analytic techniques. We consider in detail the rates, areas and the intermediate splitting scales as a function of the number of jets. In each case, we are able to characterise a general scaling behaviour characteristic for QCD processes, which we compare with results from the parton shower. The study of these observables potentially offers a very general handle in the difficult to describe regime of high-jet multiplicity.

Contents

I. Introduction	2
II. Jet Ratios from the Gen-k_T generating functional	3
A. QCD scaling limits	3
Poisson scaling	4
Staircase scaling	4
B. Inclusion of Area effects	4
Naive Phase space effects	4
Evolution equation for jet areas	5
Jet area distribution from MC	6
C. The full picture	7
III. Distribution of k_t splitting scales	8
A. General properties of splitting scales	8
One emission	9
Two emissions	9
Three emissions	11
Results for all multiplicity	11
B. Discussion of results	12
IV. Conclusions	12
A. Exponentiated form of the evolution equation	13
B. Closed solution in the staircase limit	13
References	16

*Electronic address: erik.gerwick@phys.uni-goettingen.de

†Electronic address: peter.schichtel@durham.ac.uk

I. INTRODUCTION

Quarks and gluons produced in hard scattering reactions are measured experimentally as QCD jets. Connecting hard scattering to the hadronic final state is an extended regime of parton evolution, where a large number of perturbatively well defined splittings produce particles mostly in the direction of hard partons. In some cases these also become jets so that the final state contains an appreciable number of radiated jets.

The theoretical description of radiated jets is often done via parton shower simulation [1–3]. Though very successful in some regards, there are two downsides to this. First, the parton shower is limited in formal accuracy, and second, features of the resulting distributions may be opaque from only the individual components. While we will say very little about the first point, fortunately, there is an analytic formalism, the generating functional, which allows us to address the second.

For jets defined using modern jet-finding algorithms, the generating functional was constructed first for the Durham algorithm [4]. Jet rates in this algorithm are defined in terms of a single dimensionless parameter y_{cut} . However, jets at hadron colliders are typically clustered using the generalised- k_t class of jet algorithms, for example the k_t [5, 6], anti- k_t [7] or Cambridge-Aachen [8, 9] algorithm. Here jets are defined in terms of a minimum scale E_R and radius R . A generating functional valid for these algorithms, and resumming logarithms of the type $\alpha_S \log(E/E_R) \log(1/R)$ was introduced in Ref. [10].

A useful feature of the generating functional formalism is that the multiplicity distribution can be understood to all-orders in perturbation theory. An example is the analytic solution for the average jet multiplicity [10, 11]. Similarly, the distribution differentially in the jet multiplicity can also be computed, and in the case of the Durham algorithm was in Ref. [12] for particular kinematic limits. It was found that jet rates tend to follow one of two scaling patterns [13], which are most easily understood by considering the ratio of successive jet rates, σ_{n+1}/σ_n . In the large logarithmic limit $\log(1/y_{\text{cut}}) \gg 1$, the distribution is that of a Poisson process and follows *Poisson scaling*, where $\sigma_{n+1}/\sigma_n \sim 1/(n+1)$. On the other hand, in the limit $\log(1/y_{\text{cut}}) > 1$ with $\alpha_S \log(1/y_{\text{cut}}) \ll 1$, the distribution becomes geometric $\sigma_{n+1}/\sigma_n \sim \text{constant}$, which was deemed *staircase scaling*. While some insights on scaling were given in Ref. [10] for the Gen- k_t class of jet algorithms, a complete derivation of the multiplicity distributions was not.

The motivation for this paper is to extend (and generalise) the notion of scaling in the Gen- k_t class of jet algorithms via the Gen- k_t generating functional. While the idealised scaling patterns can be worked out analytically, the distributions coming from parton shower simulation rarely follow these patterns exactly. As pointed out in Ref. [10], the discrepancy partially originates from effects which are not included in the generating functional formalism, for example kinematics and finite area effects, which disappear only in the exact limits.

For realistic jet radii ($R = 0.3 - 0.5$), a significant effect pushing jet multiplicities away from idealised staircase scaling is finite area considerations. In order to quantify this we provide in the second part of this paper a detailed analysis of the area distribution as a function of the multiplicity. Keeping with the motivation of the generating functional formalism, we attempt to describe this distribution analytically using geometric considerations and the exact collinear structure of the QCD matrix element.

In the final part of this paper, we take a step towards a more generalised notion of scaling by considering a different observable, namely the average k_t splitting scale as a function of the multiplicity. After computing some analytic results, we are able to frame the study of this observable in the context of idealised scaling patterns, one representing a perfect Poisson process, and the other for an idealised non-Abelian splitting history.

Our motivation for understanding multiplicity distributions which undergo scaling is that when these properties are sufficiently generic, they provide useful handles in difficult QCD environments. For example, while it would be very challenging to predict exact properties of a 20 (sub)-jet final state, it is much more feasible to use the shape of first principle distributions in multiplicity (comparing to say all 10 - 19 jet events) in order to constrain whether 20 (sub)-jet events are consistent with the QCD background as a whole.

A significant amount of recent progress has come from applying analytic techniques to QCD intensive observables, particularly in the context of sub-jet studies [14, 15]. In the direction of jet multiplicities and a number of other (sub)-jet properties, very recently logarithms of the type $\alpha_S \log(1/R)$ were resummed in Ref. [16], which are not considered to all-orders in the present work. Further studies on the impact of jet algorithms on resummation, and especially the appearance of logarithms in R are given in Refs. [17–19].

This paper is arranged as follows. We start off in sec.II with a brief review of the Gen- k_t generating functional and proceed to compute the multiplicity distribution. We then attempt to quantify area effects and perform a dedicated comparison to Montecarlo. In sec. III we explore the average splitting scales for different multiplicities and splitting histories. We offer some discussion on ideas in sec. IV for the applicability of this work in phenomenological studies, which are saved for future work. In the appendix we provide more details on the rate calculations performed in this work.

II. JET RATIOS FROM THE GEN- k_T GENERATING FUNCTIONAL

We start here by reviewing the Gen- k_t generating functional. A more in depth description is found in Ref. [10]. Our starting point is the fully exponentiated generating functional (see app. A) for a parton of flavour i . In terms of the opening angle $\xi = 1 - \cos \theta$ between the emitter and emitted parton, and the energy ratio of the scale evolution $e = E/E_R$, we have

$$\Phi_i(e, \xi) = u \exp \left[\int_{\xi_R}^{\xi} \frac{d\xi'}{\xi'} \int_{1/e}^1 dz \frac{\alpha_s(ze, \xi')}{2\pi} \sum_{j,k} P_{i \rightarrow jk}(z) \left(\frac{\Phi_j(e, \xi') \Phi_k(\mathcal{E}(z), \xi')}{\Phi_i(e, \xi')} - 1 \right) \right], \quad (1)$$

where $i, j, l \in \{q, \bar{q}, g\}$. The function $\mathcal{E}(z) = ze$ except for $g \rightarrow q\bar{q}$ where it is e . The one loop running coupling given by

$$\alpha_s(ze, \xi') = \frac{\pi}{b_0 \log \frac{z^2 e^2 E_R^2 \xi'}{\Lambda^2}}, \quad (2)$$

is defined in terms of the coupling at the hard scale. The -1 in the exponent of Eq. (1) defines the Sudakov form factor

$$\Delta_i(e, \xi) = \exp \left[- \int_{\xi_R}^{\xi} \frac{d\xi'}{\xi'} \int_{1/e}^1 dz \frac{\alpha_s(ze, \xi')}{2\pi} \sum_{j,k} P_{i \rightarrow jk}(z) \right], \quad (3)$$

interpreted as the no-emission probability between scales E and E_R and angular distances from the hemisphere boundary to the cut-off set by $\xi_R = 1 - \cos R$, defined in the frame of the emitter. For jet production in the entire phase space one takes $\xi = 1 - \cos(\pi/2) = 1$. If the observable is related to sub-jets inside of a larger jet of radius R_L , the correct upper boundary is $\xi = 1 - \cos(R_L)$.

Finally, the n -th jet rate is obtained by differentiating Eq. (1) with respect to the parameter u

$$\sigma_n = \sigma_0 \frac{1}{n!} \frac{d^n}{du^n} \Phi_i(e, \xi) \Big|_{u=0}. \quad (4)$$

A few notes are in order regarding the approximation of the following calculations. Analytic results are often quoted at finite order and at fixed coupling for clarity, while the full results are numerically evaluated in the plots where stated. In general, we will use the small R approximation so that $1 - \cos R \approx R^2/2$ where necessary. The logarithms we aim to control are the double-leading $\alpha_s \log(e) \log(\xi/\xi_R)$ in the exponent.

A. QCD scaling limits

Scaling patterns in jet-rates correspond to two idealised statistical distributions. The first is Poisson scaling where

$$\frac{\sigma_{n+1}}{\sigma_n} \equiv R_{\frac{n+1}{n}} = \frac{\bar{n}}{n+1}, \quad (5)$$

with \bar{n} the average number of jets. Eq. (5) is characteristic for the large logarithmically dominated regime of QCD. Staircase scaling on the other hand corresponds to the geometric (or fractal) regime of QCD radiation

$$R_{\frac{n+1}{n}} = R_0 \quad (6)$$

where R_0 is a constant. The extent to which jet distributions at ATLAS follow these patterns was studied experimentally in Ref. [20]. A correspondent pheno study can be found in [21]. Using high-precision multi-jet NLO calculations this type of behaviour was investigated in [22, 23]. In the context of BSM searches, Ref. [24] studied staircase scaling for SM backgrounds opposed to new physics decay jets, while Ref. [25] probed the extent of a staircase like distribution in sub-jet multiplicities.

We will now derive Eq. (5) and Eq. (6) in the Gen- k_t generating functional formalism.

Poisson scaling

In the double-logarithmically dominated regime of QCD we expect the rates to become a Poisson process [12, 13]. For the Durham algorithm this is achieved with a small resolution parameter y_{cut} . The generalized- k_T version depends in fact on two scale choices. The spacial jet resolution ξ_R and the allowed energy range defined by $E_R(=p_T^{\text{min}})$. In the limit $1 \geq \xi \gg \xi_R$ the integral is dominated by the $\xi' \approx \xi_R$ region. Thus we find

$$\Phi_i(e, \xi) = u \exp \left[(u-1) \int_{\xi_R}^{\xi} \frac{d\xi'}{\xi'} \int_{1/e}^1 dz \frac{\alpha_s(z, \xi')}{2\pi} \sum_{j,k} P_{i \rightarrow jk}(z) \right], \quad (7)$$

which is known to produce Poisson scaling. The limit $1/e \rightarrow 0$ depends on the structure of the splitting kernels. These have poles of the form $1/z$ which means that the $z \approx 0$ region contributes most. Therefore we also find the solution in Eq. (7). However the $g \rightarrow q\bar{q}$ splitting drops out as its splitting kernel is not divergent but goes to zero in this limit. Note that starting with several hard partons we similarly produce a Poisson process [10, 13]. However, we expect that the argument breaks down at some multiplicity around $n \approx \bar{n}$ so that $R_{\frac{n+1}{n}} < 1$, where \bar{n} is the rate parameter of the Poisson process.

Staircase scaling

From the generating functional in the Durham algorithm we know that the staircase limit is the one opposite to the large double-log limit. Formally, this regime exists when $\alpha_s \log(e) \log(1/\xi_R) \ll 1$. Thus we study Eq. (1) in the $e \rightarrow 1$ limit. To work within the framework of the generating functional, which resums large logarithms, we therefore need $\log(1/\xi_R) \gg 1$. We also focus on the pure Yang-Mills case for simplicity, though the arguments are more general. We Taylor expand the integrand of the generating functional around $z_0 \approx 1$ but with $z < 1$. Thus, we can write Eq. (1) in the form

$$\Phi_g(e, \xi) = u \exp \left[\int_{\xi_R}^{\xi} \frac{d\xi'}{\xi'} \int_{1/e}^1 dz \frac{\alpha_s(z, \xi')}{2\pi} P_{g \rightarrow gg}(z) \left(\Phi_g(e, \xi') + \sum_{n=1}^{\infty} \frac{(e(z-1))^n}{n!} \frac{d^n \Phi_g(e, \xi')}{de^n} - 1 \right) \right]. \quad (8)$$

Taking only the leading behavior of the $n = 1$ term into account we are able to find a closed solution for Φ_g (see app. B)

$$\Phi_g(e, \xi) = \frac{1}{1 + \frac{(1-u)}{u\Delta_g(e, \xi)} - (u-1)\chi(e, \xi)} \approx \frac{1}{1 + \frac{(1-u)}{u\Delta_g(e, \xi)}} + \frac{(u-1)\chi(e, \xi)}{(1 + \frac{(1-u)}{u\Delta_g(e, \xi)})^2}. \quad (9)$$

The function $\chi(e, \xi)$ is given in the app. B. By taking successive u derivatives we find that the jet ratios are

$$R_{\frac{n+1}{n}} = (1 - \Delta_g(e, \xi)) \left[1 + \left(\frac{(1 - \Delta_g(e, \xi))^3}{\chi(e, \xi)\Delta_g^2(e, \xi)} - \frac{1}{\Delta_g(e, \xi)} - 2 + (n+1) \right)^{-1} \right]. \quad (10)$$

We can check the large n limit of this formula by taking the resolved limit $\alpha_s L^2 \ll 1$ in which case the term in brackets behaves like $1 + (1/n)$. This is in agreement with the fitted form of the resolved coefficients as derived in Ref. [26].

The result in Eq. (10) is very similar to the one found using the Durham algorithm in [12]. At high multiplicity the ratios converge to constant staircase scaling. However, we find a very interesting additional feature: a staircase breaking term indicating that pure staircase scaling is an asymptotic feature in n . The breaking term enhances the low multiplicity ratios with respect to pure staircase scaling. Note that, as mentioned above, formally we need small ξ_R . For finite values we expect ξ_R dependent deviations from Eq. (10).

B. Inclusion of Area effects

Naive Phase space effects

Since we will be studying the high multiplicity behaviour in detail, an additional effect we would like to include is the simple depletion of geometric area as the jet multiplicity increases. The assumptions we make are that the

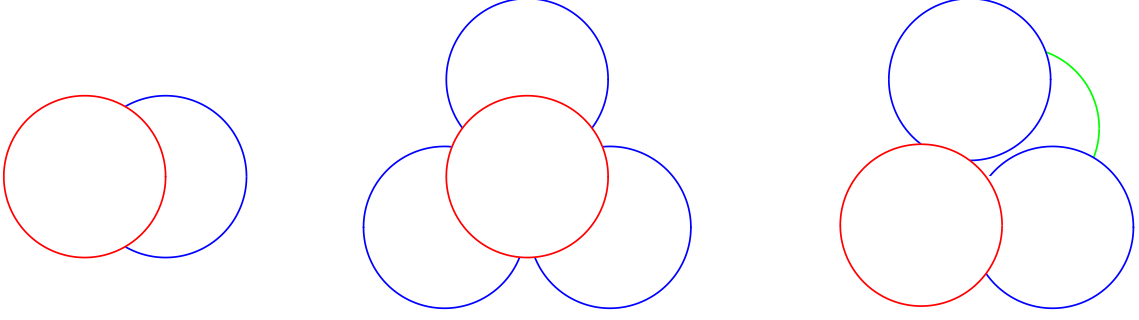


Figure 1: Left: The harder jet (red) defines a disk like object, while the softer jet (blue) is partially covered. This defines Eq. (15). Middle: Flower like configuration of collinear jets as included in our analytic approach. Right: Non-trivial overlap configuration demonstrating that very small jet areas are possible (green).

geometric area of each jet is approximated by πR^2 , and that the rates decrease proportional to the available area uniformly. This uniquely amounts to an overall suppression factor at the level of the ratios

$$\phi_{\text{ps}} \equiv \frac{\phi(n+1)}{\phi(n)} = \frac{1 - (n+1)R^2/4}{1 - nR^2/4} \quad (11)$$

so that the expected behaviour is described by $R_{\frac{n+1}{n}} \rightarrow \phi_{\text{ps}} R_{\frac{n+1}{n}}$. At some n the numerator will reach 0 signifying that naively, the entire phase is saturated, and the ratios go to 0. Although we will later see that this is an insufficient approximation it is very useful for some order of magnitude estimates. We investigate the exact phase space effects in detail in the next paragraphs.

Evolution equation for jet areas

In this section, we study how the distribution of jet areas $\rho(A)$ evolves with higher multiplicity. Jet areas in the generalised- k_t class of jet algorithms was studied first in Ref. [27], with many techniques implemented in Ref [28]. It is a well known feature of the anti- k_T algorithm that it produces circular jets [7]. Furthermore, only with this algorithm is the area observable IR safe. We therefore orient this section towards specifically anti- k_T .

Due to the collinear nature of QCD jets quite often overlap, as depict schematically in the left panel of Fig. 1. This fact causes the jet areas to follow a non-trivial distribution in multiplicity. We assert that for a fixed number of jets n the distribution of areas can be parameterised as

$$\rho_n(A) = \bar{\rho}_n \delta(A - \bar{A}^{\text{full}}) + (1 - \bar{\rho}_n) \rho_n^{\text{overlap}}(A), \quad (12)$$

where $\bar{A}^{\text{full}} = 2\pi(1 - \cos(R)) \approx \pi R^2$ is just the geometric size of a jet of radius R projected on the sphere and $\bar{\rho}_n$ denotes how many full area jets we expect, which encodes one part of the non-trivial evolution. The other non-trivial part is described by $\rho_n^{\text{overlap}}(A)$, which is a result of how the jets overlap.

What can we learn about these two quantities? Let us start with the simplest case $n = 2$, where $\bar{\rho}_2 = 1$. Therefore, we find trivially

$$\rho_2(A) = \delta(A - \bar{A}^{\text{full}}). \quad (13)$$

From the collinear approximation we know that jets are distributed due to $d\xi/\xi$. Assuming fixed coupling and noting that the full area jet, which is already present, yields a lower bound for the integration we find

$$\bar{\rho}_3 = \frac{\log(4) - 2 \log(\csc^2(R))}{\log(1 - \cos(R))}. \quad (14)$$

To find $\rho_n^{\text{overlap}}(A)$ we need to study how the geometrical overlap of two jets works. For $n = 3$ we can construct

that the non-covered part of the area of a collinear jet which is partially covered by a full area jet is

$$\begin{aligned} A^{\text{not-covered}}(\xi) &= R \left(\sqrt{2\xi - \frac{\xi^2}{2R^2}} + 2R \arcsin \left(\sqrt{\frac{\xi}{2R^2}} \right) \right) \\ &\approx \frac{1}{3}\pi R^2 - \sqrt{\frac{1}{3}} \left(2R^2 - R\sqrt{32\xi + \xi} \right). \end{aligned} \quad (15)$$

Noting that Eq. (15) implies a lower bound of

$$\bar{A}_3^{\min} = \frac{1}{6} \left(3\sqrt{3} + 2\pi \right) R^2 \quad (16)$$

for the minimal jet size and using again that jets are distributed due to $d\xi/\xi$ we compute

$$\begin{aligned} \rho_3^{\text{overlap}}(A) &= \frac{6\Theta((\bar{A}^{\text{full}} - A)(A - \bar{A}_3^{\min}))}{6A - 2(6\sqrt{3} + \pi)R^2 + 4\sqrt{2}3^{\frac{1}{4}}R\sqrt{-3A + (6\sqrt{3} + \pi)R^2}} \\ &\times \frac{1}{\log\left(4 + \frac{4\pi}{\sqrt{3}}\right) - 2\operatorname{arccoth}\left(\frac{23^{\frac{1}{4}}}{\sqrt{\sqrt{27}-\pi}}\right)}. \end{aligned} \quad (17)$$

In the collinear region this equation is exact. One might be concerned that the transition to the δ -function for full area jets is not smooth. Indeed, if we expand Eq. (15) around $\xi = (2R)^2/2$ instead of $R^2/2$ we find a power behaviour with exponent $3/2$. This causes the measure of the inverse function to diverge. The result is a divergent, yet integrable, behaviour of $\rho_3^{\text{overlap}}(A \approx \pi R^2)$. We do not try to merge the two behaviours here, because this region is shielded by finite resolution effects as we will see in the following section.

For higher jet multiplicities the situation gets more complicated. Schematically moving from n to $n+1$ jets is described by

$$\rho_{n+1}(A) = \frac{1}{n+1} \left(n\rho_n(A) + \bar{\rho}_{n+1,1} \delta(A - \bar{A}^{\text{full}}) + (1 - \bar{\rho}_{n+1,1}) \rho_{n+1,1}(A) \right). \quad (18)$$

Here $1 - \bar{\rho}_{n+1,1}$ is how likely it is to add another covered jet. With $\rho_{n+1,1}(A)$ we describe how the area of this jet is distributed. For lower multiplicities we might hope to describe both with their ρ_3 equivalents. Deviations from this assumption come from configurations where the additional jet overlaps with more than one of the previous ones, which is a configuration we cannot easily describe with naive geometrical considerations. Of course, from a certain multiplicity on these configurations will dominate. One of the features that result from this fact is that we will not have a sharp drop at \bar{A}_3^{\min} anymore, but a smoother distribution allowing all possible values for A . In order to estimate at which value n for the jet multiplicity this will become significant, let us perform some counting gymnastics for $R = 0.5$. Some leading jets could cover three to four other jets without them touching each other. Such a configuration would look like a kids version of a flower, see center panel of Fig. 1. For tighter configurations we are forced to produce non-trivial overlaps as shown in the right panel of Fig. 1. Therefore, we should see a non-trivial structure from starting at 5 or 6 jets. However, choosing this exact configuration is rather unlikely, if we remember that there is a chance $\mathcal{O}(\bar{\rho}_3)$ to produce full area jets instead. From the naive phase space considerations, see Eq. (11), we know that around $n = 15$ the non-overlap picture must break down. This means that between $n = 6$ and $n = 15$ the non-trivial structure becomes significant. The simplest guess would be that we are able to describe jet area distributions with successive usage of ρ_3 up to $n \approx 10$. Using Eq. (18) we see that we have $\langle A_n \rangle < \bar{A}^{\text{full}}$. This leads to more jets than naively expected. However, generalizing configurations like in the right panel of Fig. 1 it is clear that this still under-estimates the maximum number of possible jets. The end point of the spectrum is non-trivial and we will study it in more detail in the next section.

Jet area distribution from MC

In this paragraph we would like to compare our analytic considerations with a parton shower simulation (specifically the CS shower in SHERPA [29, 30]) in order to assess non-trivial behaviour at high multiplicities. For this purpose it is instructive to recall how jet areas are measured. In the Fastjet implementation [27, 28] jet areas can be computed in one of three ways, active area, Passive areas and Veronoi areas. We use the active area option of Fastjet, but implement spherical coordinates. The active area option divides the (η, ϕ) -plane into cells of fixed size and places a ghost particle (almost vanishing p_T) randomly in each cell. If the ghost happens to end up in a certain jet, we count the cell's size towards that jet's active area.

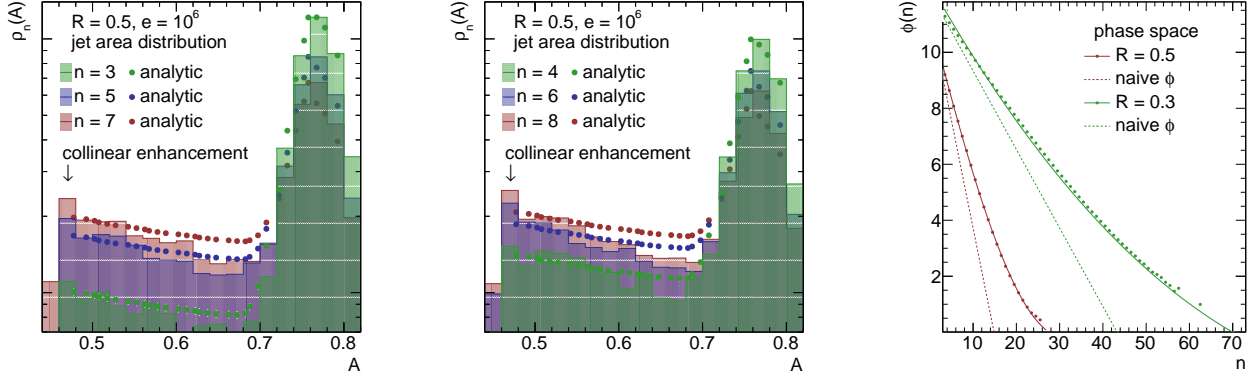


Figure 2: Left and middle panel: Comparison of $\rho_n(A)$ from Eq. (12) (dots) with MC data for $R = 0.5$ from $n = 3$ up to $n = 8$ (shaded)xs. Right panel: Comparison of the actual phase space computed via $\langle A(n) \rangle$ (dots) with the naive formula Eq. (11) (dashed) and a polynomial of third order (solid) for $R = 0.5$ (red) and $R = 0.3$ (green).

We study jets of size $R = 0.5$ and $R = 0.3$. We will also consider jets of size $R = 0.1$ in the next section. However, we will not study them here, simply because even from naive phase space considerations we only expect effects after $n \approx 400$. Let us estimate what we expect for the area measurement. To compensate for the flat approximation we use an effective radius which corresponds to an area of \bar{A}^{full} . Using a random sample of ghosts and jet midpoints we expect that the area is normal distributed around $A = 0.77$ with a width of $\Delta A = 0.02$ for $R = 0.5$ jets and $A = 0.28$ with a width of $\Delta A = 0.016$ for $R = 0.3$ jets. The variance ΔA is driven by the cell size and the actual jet radius.

We compare these the analytic formula from the previous section with MC data, using the numbers in the previous paragraph. For the average jet size and width we find good agreement. However, the actual width is slightly larger due to spherical geometry. In Fig. 2 we show the non-trivial area distributions for the exclusive $n = 3$ up to $n = 8$ cases for $R = 0.5$. We find good agreement for the three jet case with Eq. (17). For higher multiplicities we fix the overall normalisation of Eq. (18) to produce the same maximum height as the peak at \bar{A}^{full} in the MC*. We observe that we describe the area overlap distribution accurately for low multiplicities. As expected, for higher multiplicities $n \approx \mathcal{O}(10)$ our description breaks down. In the right panel of Fig. 2 we show the phase space for all multiplicities for $R = 0.5$ and $R = 0.3$ together with Eq. (11) for the naive phase space expectation. For that purpose we compute $\phi(n) = 4\pi - \langle A(n) \rangle$. The true phase space follows a polynomial of third order, where the linear part is driven by Eq. (17). The higher order terms encode the non-trivial overlap configurations not accessible with our ansatz.

C. The full picture

Now we would like to put all the different parts together, and see to what extent we can understand the effects driving the jet ratio distribution. In order to maximise statistics we choose a very large energy hierarchy $E = 10^6$ GeV and $E_R = 1$ GeV, with jets clustered in the anti- k_t jet algorithm.

We expect low multiplicities to be described by a Poisson process, defined by $\bar{n} = |\log(\Delta(e, \xi))|$ [12]. In addition the average number of jets is rather sensitive to the chosen coupling at the hard scale, which we take as $\alpha_S(E) = 0.076$. We show a comparison in Tab. I and find good agreement in the double logarithmic region. Finite R effects are expected from the generating functional as well as the fact that very low jet multiplicities are matched to the matrix element. At multiplicities $n \gg \bar{n}$ we have Eq. (10). In between, there is a somewhat awkward intermediate regime which we cannot say much about. However, for realistic jet radii $\mathcal{O}(0.5)$ this regime is very small. For high multiplicities the pure staircase scaling is suppressed by phase space effects driven by depletion of available area as shown in Fig. 2. Note, however, that for finite R we expect additional effects, which are not included in our approach. These should vanish as $R \rightarrow 0$.

To demonstrate these claims we show exclusive jet cross section ratios for $R = 0.5$, $R = 0.3$ and $R = 0.1$ with $e = 10^6$ in Fig. 3. The green line corresponds to the Poisson hypothesis, while the yellow dashed line

* This is necessary because we neglect any possibility of non-trivial overlap for our analytic ansatz. This directly leads to a different normalization for the two quantities we wish to compare, which we compensate with this prescription.

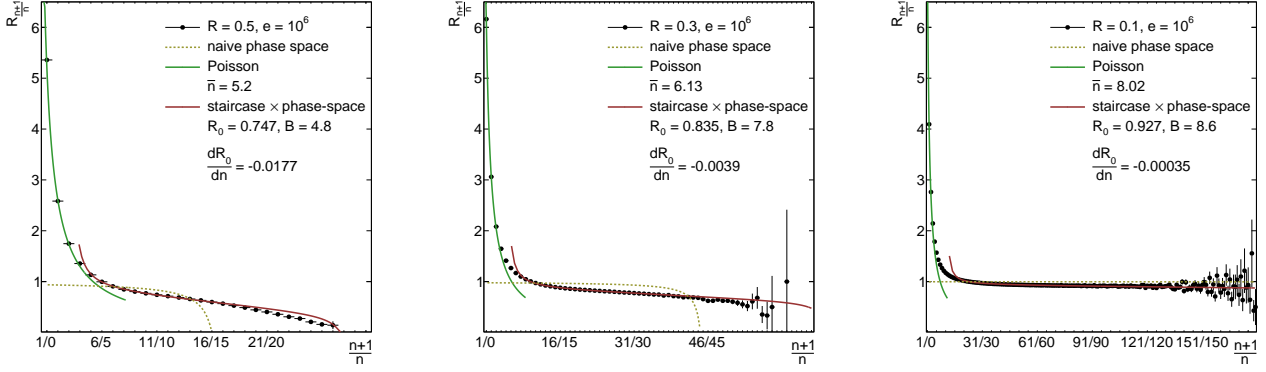


Figure 3: Exclusive jet cross section ratios for $R = 0.5$ (left), $R = 0.3$ (middle), and $R = 0.1$ (right). The red solid lines show the staircase scaling hypothesis times phase space. An additional term $dR_0/dn(n+1)$ is fitted to quantify the deviation from Eq. (10). In addition we show the naive phase space assumption (dashed yellow) and the Poisson scaling hypothesis (green).

specifies Eq. (11). The red line is characterized by (see Eq. (10)),

$$R_{\frac{n+1}{n}} = \left(R_0 \left[1 + \frac{1}{B + (n+1)} \right] + \frac{dR_0}{dn}(n+1) \right) \times \frac{\phi(n+1)}{\phi(n)}, \quad (19)$$

where the numerical values for R_0 , B , and dR_0/dn are given in the plots. We take the phase space factor from the polynomial in the previous section. Note, while there is a deeper connection between R_0 and B , as well as their exact position Eq. (19) that is not true for the term dR_0/dn . This term is purely phenomenological and we introduce it to describe an additional small tilt we observe for jet cross section ratios. Therefore it quantifies the discrepancy from our staircase scaling formula. Recall that we formally need small R for the derivation of Eq. (10). In Tab. I we show that the fitted deviation dR_0/dn is small and indeed vanishes as $R \rightarrow 0$.

III. DISTRIBUTION OF k_t SPLITTING SCALES

Intermediate splitting scales in multi-jet events are useful primarily due to the close correspondence with parton shower splitting variables. They have been measured extensively in collider experiments, in e^+e^- [31] and more recently at the LHC in the context of ME-PS matching [32]. As our emphasis in this work is scaling features especially at high-multiplicity, we explore in this section the distribution of splitting scales as a function of multiplicity very much in analogy to the jet rates. The hope is that the resulting distribution could provide generic first principle handles on QCD showered events, in connection to what (Staircase/Poisson) scaling behaviour provides for the rates.

A. General properties of splitting scales

In order to discuss splitting scales as derived from the Gen- k_t generating functional, we first note that emissions are separately ordered in energy and angle in this framework, and not explicitly in k_t . We can enforce k_t ordering by hand, with a Θ function for example, bearing in mind that $zE\sqrt{\xi/\xi_R}$ corresponds to the usual definition of the splitting scale for a k_t jet algorithm [5, 33] in the small- R limit.

Our notation for the splitting scales is that $\langle k_t^{(i,j)} \rangle$ is the j hardest emission in k_t in a i jet splitting history, so that we will always speak of the distribution or ordered scales for an exclusive event. Now let us discuss briefly what we expect the resulting distributions to convey regarding the nature of QCD showering.

First of all, since we know that the Pseudo-abelian limit where all emissions are primary corresponds to a Poisson process, the average intermediate scale $\langle k_t^{(i)} \rangle$ is independent of the multiplicity. This property is true both at the inclusive and exclusive level, indicating that it persists for fixed-order and resummed calculations. A way to see that this is that the splitting scales correspond to the inter-arrival times of the Poisson process, which are themselves memory-less.

R	0.5	0.4	0.3	0.2	0.1	0.09	0.08	0.07	0.06	0.05	0.04
$ \log(\Delta(e, \xi)) $	3.2	3.8	4.7	6.0	8.3	8.7	9.1	9.6	10.1	10.7	11.5
\bar{n}	5.2	5.6	6.1	6.9	8.0	8.9	9.2	9.7	10.0	10.4	11.1
R_0	0.75	0.79	0.82	0.857	0.921	0.934	0.937	0.938	0.941	0.945	0.953
$ \frac{dR_0}{dn} $	0.018	0.010	0.0039	0.0011	0.00035	0.0003	0.0002	0.0002	0.00013	0.00011	0.000074

Table I: We summarize the relevant quantities for studying deviations as a function of R . The first two rows show the expected and observed Poisson parameter \bar{n} . The last two rows contain the numerical value for R_0 compared to dR_0/dn .

The constant average splitting scales in the pseudo-Abelian limit is similar in origin to the idealised $\sigma_{n+1}/\sigma_n = 1/(n+1)$ behaviour in the rates for a Poisson process. In QCD, once we allow for correlated emissions, the picture will change, and we may ask the question whether there is some notion of a correlated emission dominated phase imprinted in the splitting scales. As we will see, our computations indicate that there is some evidence for behaviour like this.

One emission

We start with the single emission splitting scale. This is the average value of $Ez\sqrt{\xi'/\xi_R}$ of the probability function $(z\xi')^{-1}\alpha_S(zE\xi')\Delta(z, \xi')$ in the plane defined by boundaries $\xi' \in [\xi_R, \xi]$ and $z \in [1/e, 1]$. The expectation value of the splitting scale $\langle k_t \rangle = zE\sqrt{\xi'/\xi_R}$ inside the integral is

$$\langle k_t^{(1,1)} \rangle = \frac{E \int_{\xi_R}^{\xi} \frac{d\xi'}{\sqrt{\xi'}} \int_{1/e}^1 dz z P(z) \alpha_S(zE\xi') \Delta(z, \xi')}{\sqrt{\xi_R} \int_{\xi_R}^{\xi} \frac{d\xi'}{\xi'} \int_{1/e}^1 dz P(z) \alpha_S(zE\xi') \Delta(z, \xi')} \quad (20)$$

which keeping only the most singular contributions gives the leading logarithms

$$\langle k_t^{(1,1)} \rangle = \frac{2(E - E_R)(\sqrt{\xi/\xi_R} - 1)}{\log(E/E_R) \log(\sqrt{\xi/\xi_R})} + \mathcal{O}(\alpha_S) \quad (21)$$

The expression in Eq. (20) is well approximated by Eq. (21) for all but large energy scale ratios $\mathcal{O}(10^6)$, or extremely small jet radii. This effect can be seen by noting that the inclusion of higher order terms in the Sudakov affects the numerator and denominator of Eq. (20) in the same direction, thus leaving a diminished residual dependence. Furthermore, there is no leading order dependence on α_S , a consequence is that running coupling effects are pushed to third-order.

Two emissions

In order to compute the average splitting scales for two emission we distinguish between $k_t^{(21)}$ and $k_t^{(22)}$ such that we have $k_t^{(21)} > k_t^{(22)}$. Furthermore, at this multiplicity there are two splitting histories which we denote correlated and uncorrelated



We write the two primary emission rate in a k_t ordered way such that the contribution to lowest order is

$$P_2^{(\text{uncorr})} = \frac{1}{2} \frac{\alpha_S C_F^2}{\pi} \left(\int_{\xi_R}^{\xi} \frac{d\xi'}{\xi'} \int_{E_R/E}^1 dz P(z) \int_{\xi_R}^{\xi} \frac{d\xi''}{\xi''} \int_{E_R/E}^1 dz' P(z') \Theta(z' \sqrt{\xi''} < z \sqrt{\xi'}) \right. \\ \left. + \int_{\xi_R}^{\xi} \frac{d\xi'}{\xi'} \int_{E_R/E}^1 dz P(z) \int_{\xi'}^{\xi} \frac{d\xi''}{\xi''} \int_{E_R/E}^1 dz' P(z') \Theta(z' \sqrt{\xi''} \geq z \sqrt{\xi'}) \right) \quad (22)$$

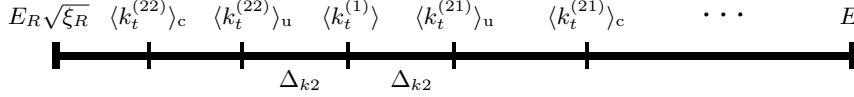


Figure 4: Schematic distribution for the average splitting scales $\langle k_t \rangle$ for 1 and 2 gluon emissions, in the latter case split in terms of the correlated and uncorrelated component. The scale E represents the initial energy of the emitting parton.

which gives the well known result $(1/2!)P_1^2$ for the rate. The contributions in Eq. (22) arise when the second emission in the integral is smaller or larger in k_t respectively. For correlated emissions, due to ordering only the first term is present.

The Θ function provides a complicated phase space constraint, preventing a simple analytic evaluation. In the inclusive k_t (Durham) algorithm, the k_t regions would factorise, and we would expect a much simpler final result. However, in this work we have stayed as close as possible to the more typically used jet variable E_R and ξ_R . Despite this complication, we can extract the average splitting scale of the harder emission

$$\begin{aligned} \langle k_t^{(21)} \rangle_{\text{uncorr.}} = & \frac{E}{P_2^{(\text{uncorr})}} \frac{\alpha_S C_F^2}{2\pi} \left(\int_{\xi_R}^{\xi} \frac{d\xi'}{\xi'} \int_{E_R/E}^1 dz P(z) z \sqrt{\xi'} \int_{\xi_R}^{\xi} \frac{d\xi''}{\xi''} \int_{E_R/E}^1 dz' P(z') \Theta(z' \sqrt{\xi''} < z \sqrt{\xi'}) \right. \\ & \left. + \int_{\xi_R}^{\xi} \frac{d\xi'}{\xi'} \int_{E_R/E}^1 dz P(z) \int_{\xi'}^{\xi} \frac{d\xi''}{\xi''} \int_{E_R/E}^1 dz' P(z') z' \sqrt{\xi''} \Theta(z' \sqrt{\xi''} > z \sqrt{\xi'}) \right), \end{aligned} \quad (23)$$

and similarly for the softer

$$\begin{aligned} \langle k_t^{(22)} \rangle_{\text{uncorr.}} = & \frac{E}{P_2^{(\text{uncorr})}} \frac{\alpha_S C_F^2}{2\pi} \left(\int_{\xi_R}^{\xi} \frac{d\xi'}{\xi'} \int_{E_R/E}^1 dz P(z) \int_{\xi_R}^{\xi} \frac{d\xi''}{\xi''} \int_{E_R/E}^1 dz' P(z') z' \sqrt{\xi''} \Theta(z' \sqrt{\xi''} < z \sqrt{\xi'}) \right. \\ & \left. + \int_{\xi_R}^{\xi} \frac{d\xi'}{\xi'} \int_{E_R/E}^1 dz P(z) z \sqrt{\xi'} \int_{\xi'}^{\xi} \frac{d\xi''}{\xi''} \int_{E_R/E}^1 dz' P(z') \Theta(z' \sqrt{\xi''} > z \sqrt{\xi'}) \right), \end{aligned} \quad (24)$$

We can evaluate these analytically by taking only the energy ordering enforced by the Θ functions, which differs from the full result by sub-leading terms, and demonstrates the same essential behaviour. In this case we obtain

$$\begin{aligned} \langle k_t^{(21)} \rangle_{\text{uncorr.}} &= \frac{4}{\log^2(E/E_R) \log(\xi/\xi_R)} \left(E_R - E + E \log \left[\frac{E}{E_R} \right] \right) (\sqrt{\xi/\xi_R} - 1), \\ \langle k_t^{(22)} \rangle_{\text{uncorr.}} &= \langle k_t^{(21)} \rangle_{\text{uncorr.}} \frac{E - E_R + E_R \log(E_R/E)}{E_R - E + E \log(E/E_R)}. \end{aligned}$$

This also gives the symmetric fluctuation about the uncorrelated average in the individual components as $\Delta_{k2} = \langle k_t^{(1)} \rangle - \langle k_t^{(22)} \rangle$ (see Fig. 4).

Turning to the correlated emissions, due to the energy and angular ordering in the correlated emission, this contribution is easily evaluated and gives

$$\begin{aligned} \langle k_t^{(21)} \rangle_{\text{corr.}} &= \frac{E}{P_2'} \frac{\alpha_S C_F C_A}{\pi} \int_{\xi_R}^{\xi} \frac{d\xi'}{\xi'} \int_{E_R/E}^1 dz P(z) z \sqrt{\xi'} \int_{\xi_R}^{\xi'} \frac{d\xi''}{\xi''} \int_{E_R/E}^z dz' P(z') \\ &= \frac{4}{\log^2(E/E_R) \log^2(\xi/\xi_R)} \left(E_R - E + E_R \log \left[\frac{E}{E_R} \right] \right) \left(2 - 2\sqrt{\xi/\xi_R} + \sqrt{\xi/\xi_R} \log \left[\frac{\xi}{\xi_R} \right] \right), \end{aligned} \quad (25)$$

and

$$\begin{aligned} \langle k_t^{(22)} \rangle_{\text{corr.}} &= E \frac{E}{P_2'} \frac{\alpha_S C_A C_F}{\pi} \int_{\xi_R}^{\xi} \frac{d\xi'}{\xi'} \int_{E_R/E}^1 dz P(z) \int_{\xi_R}^{\xi'} \frac{d\xi''}{\xi''} \int_{E_R/E}^z dz' P(z') z' \sqrt{\xi''} \\ &= \frac{4}{\log^2(E/E_R) \log^2(\xi/\xi_R)} \left(E - E_R + E_R \log \left[\frac{E_R}{E} \right] \right) \left(2\sqrt{\xi/\xi_R} - 2 - 2 \log \left[\frac{\xi}{\xi_R} \right] \right). \end{aligned} \quad (26)$$

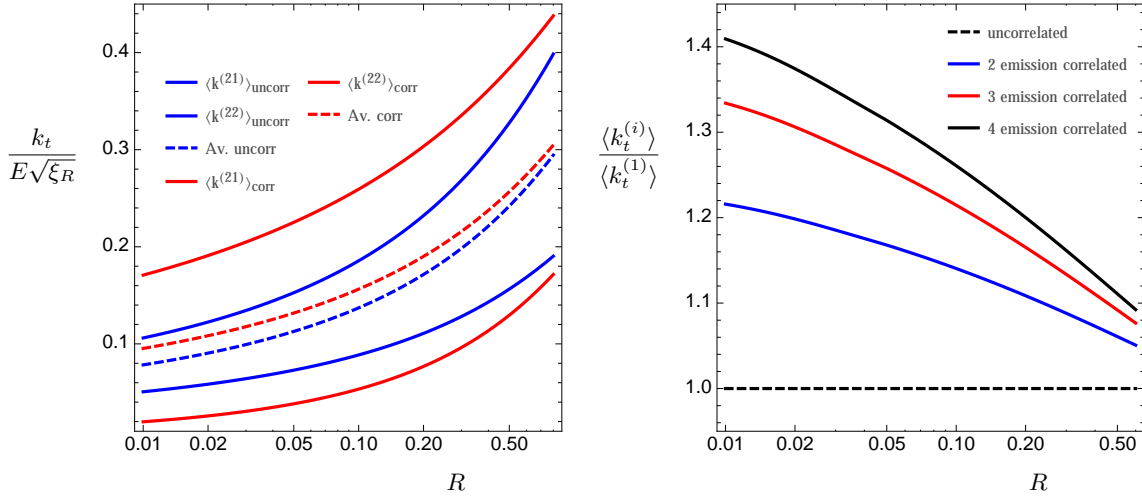


Figure 5: (Left) Distribution for the scaled average splitting scale $\langle k_t \rangle / E$ for 2 gluon emission as a function of R , divided among the correlated and uncorrelated components. (Right) Ratio of the average splitting scale per jet for 1,2,3,4 emissions for the all correlated contribution.

Three emissions

For the 3 emission component we proceed as before though we have 4 splitting histories, which differ by their sequence of primary (P) and secondary (S) emissions. They can be written PPP , PPS , PSP and PSS where

$$PPP : \text{diagram} \quad PPS : \text{diagram} \quad PSP : \text{diagram} \quad PSS : \text{diagram}$$

The diagrams represent the splitting histories: PPP (three primary emissions), PPS (two primary, one secondary), PSP (one primary, two secondary), and PSS (one primary, two secondary in a different order).

First we have the normal Poisson relation that $\langle k_t^{(32)} \rangle_{PPP} = \langle k_t^{(1)} \rangle$ and $\langle k_t^{(31)} \rangle_{PPP} - \langle k_t^{(32)} \rangle_{PPP} = \langle k_t^{(32)} \rangle_{PPP} - \langle k_t^{(33)} \rangle_{PPP}$ in the all correlated part. We give here only the analytic form of the PSS (or fully correlated) contribution for the hardest emission

$$\begin{aligned} \langle k_t^{(31)} \rangle_{PSS} = & \frac{288}{\log^3(E/E_R) \log^3(\xi/\xi_R) E_R} \left(E \log \left(\frac{E}{E_R} \right) (E_R - 4E) \log \left(\frac{E}{E_R} \right) + 4E - 2E_R - 2E + 2E_R \right) \\ & \left(\sqrt{\xi/\xi_R} \left(\frac{1}{8} \log^2 \left(\frac{\xi}{\xi_R} \right) - \frac{1}{2} \log \left(\frac{\xi}{\xi_R} \right) + 1 \right) - 1 \right). \end{aligned} \quad (27)$$

As with the 2 correlated emissions the average of the splitting scale for 3 correlated emissions $(\langle k_t^{(31)} \rangle + \langle k_t^{(32)} \rangle + \langle k_t^{(33)} \rangle_{PSS})/3$ is larger than the single emission result $\langle k_t^{(1)} \rangle$.

Results for all multiplicity

In principle we can repeat the procedure for an arbitrary splitting history, and thus for all multiplicities. However, the current method is extremely tedious. Instead we will focus only on the all correlated contribution since it has the largest deviation from a Poissonian splitting distribution for a fixed multiplicity. In other words, all average splitting scales for an arbitrary history fall somewhere between the all correlated and the all primary results. The leading contribution to the all correlated average at fixed coupling is obtained from the limit

$$\langle k_t^{(\infty)} \rangle = \lim_{j \rightarrow \infty} \frac{\left(E \prod_{k=1}^j \int_{\xi_R}^{\xi_{k+1}} \frac{d\xi_k}{\xi_k} \int_{E_R/E}^{z_{k+1}} \frac{dz_k}{z_k} \right) \sum_{k=1}^j z_k \sqrt{\xi_k/\xi_R}}{j \left(\prod_{k=1}^j \int_{\xi_R}^{\xi_{k+1}} \frac{d\xi_k}{\xi_k} \int_{E_R/E}^{z_{k+1}} \frac{dz_k}{z_k} \right)} \quad (28)$$

which for the general case we are not able to take analytically. However a trick at our disposal is the small R limit, where the small R behaviour of the correlated plots in Fig. 5 suggests that this converges to a finite non-zero value. We find the following series representation for Eq. (28) divided by the single emission average

in that case,

$$\frac{\langle k_t^{(\infty)} \rangle}{\langle k_t^{(1)} \rangle} = \lim_{j \rightarrow \infty} \frac{(j+1)!}{(1-e) \log^j(e)} \left(1 - e + \sum_{k=1}^j \frac{\log^k(e)}{k!} \right) + \mathcal{O}(1/\log(1/R)). \quad (29)$$

Substituting the asymptotic Stirling approximation for the the sum in brackets, we find the surprisingly simple expression

$$\begin{aligned} \frac{\langle k_t^{(\infty)} \rangle}{\langle k_t^{(1)} \rangle} &= \lim_{j \rightarrow \infty} \frac{(j+1)!}{(1-e) \log^j(e)} \left(-\frac{\log^{j+1}(e)}{j^{j+1} \exp(-j) \sqrt{2\pi j}} \right) + \mathcal{O}(1/\log(1/R)) \\ &= \frac{\log(e)}{e-1} \end{aligned} \quad (30)$$

This in turn implies the small R behaviour of the correlated emission scale $\langle k_t^{(\infty)} \rangle = E/\log(\sqrt{2}/R)$, which can be compared to the pure Poisson single emission result at small- R , given from Eq. (21) by $\langle k_t^{(1)} \rangle = E(1-e)/(\log(e) \log(\sqrt{2}/R))$. Note that these expression very nicely enforce the intuitive fact that $\langle k_t^{(1)} \rangle < \langle k_t^{(\infty)} \rangle$ for all values of E_R and R , even though the latter does not depend on E_R explicitly. The correlated emission component, via the precise cancellation in Eq. (29) in the $R \ll 1$ limit loses memory of the scale E_R .

B. Discussion of results

We summarise the findings of our calculations for the average k_t splitting scales. For arbitrary multiplicity, the all uncorrelated contribution produces a Poisson process where the average splitting scale is a constant for all n . The largest deviation from a Poisson process is the all correlated emission component, which puts this observable in a similar class as the average jet multiplicity [34], in the sense of being maximally sensitive to correlated emissions. We were able to show that at leading order in the small- R limit, the all-correlated contribution converges to a constant. Of course physically it must converge to something since it is bounded from above and below, but what is surprising is the simple analytic result in this case. We speculate that for the generic leading logarithmic QCD behaviour of the average jet multiplicity it converges to a constant

$$\lim_{n \rightarrow \infty} \frac{\langle \bar{k}_t^{(n)} \rangle}{\langle \bar{k}_t^{(1)} \rangle} = 1 + c, \quad (31)$$

with $c > 0$. Whether or not this behaviour is realised in realistic QCD environments, for example in the internal dynamics of large jets, is a question which can be answered with a dedicated phenomenological study.

IV. CONCLUSIONS

In this paper we have studied jets at high multiplicity. Using the Gen- k_t generating functional, we found that scaling patterns in the rates emerged in particular limits in energy, angle and multiplicity. We speculated that finite area effects are responsible for depressing the geometric scaling in the jet ratios and attempted to quantify the area distribution as accurately as possible. Using these considerations we managed to describe the area distribution coming from a parton shower simulation. The area depletion effect on the rates on the other hand, was found to be significant but too small to explain the tilt in the n -jets distribution at finite R , indicating that additional terms are necessary to fully, analytically describe the parton shower generated n_{jet} distribution. In the second part of this paper we looked into a new observable from the perspective of scaling and found that it could also be shown, formally at least, to produce a type of universal behaviour at high multiplicity, which emerges due to the dominance of correlated emissions in this limit.

We did not undertake a detailed phenomenological study in this work. However, one can imagine that sub-jet multiplicities with an energy cut sufficiently small might also show an extended regime of generalised types of scaling. Furthermore, the distribution of jet areas can be extended to regimes where this quantity is useful, for example assessing pile-up contributions and non-perturbative effects. Finally, we speculate that additional distributions and observables may show surprising emergent properties at high multiplicity, which could lead to experimental and theoretical progress.

Acknowledgments

We thank Steffen Schumann for helpful comments on the draft. PS acknowledges support from the european Union as part of the FP7 Marie Curie Initial Training Network MCnet ITN (PITN-GA-2012-315877) and the IMPRS for Precision Tests of Fundamental Symmetries. EG acknowledges support by the Bundesministerium für Bildung und Forschung under contract 05H2012.

Appendix A: Exponentiated form of the evolution equation

We start from Eq. (4.7) in [10], which is

$$\begin{aligned}\Phi_q(E, \xi) &= u + \int_{\xi_R}^{\xi} \frac{d\xi'}{\xi'} \int_{E_R/E}^1 dz \frac{\alpha_s(k_T^2)}{2\pi} P_{q \rightarrow qg} \Phi_q(E, \xi') (\Phi_g(zE, \xi') - 1) \\ \Phi_g(E, \xi) &= u + \int_{\xi_R}^{\xi} \frac{d\xi'}{\xi'} \int_{E_R/E}^1 dz \frac{\alpha_s(k_T^2)}{2\pi} [P_{g \rightarrow gg} \Phi_g(E, \xi') (\Phi_g(zE, \xi') - 1) \\ &\quad + P_{g \rightarrow q\bar{q}} (\Phi_q^2(E, \xi') - \Phi_g(E, \xi'))].\end{aligned}\tag{A1}$$

Using $e = E/E_R$ we rewrite it in the form

$$\Phi_i(e, \xi) = u + \int_{\xi_R}^{\xi} \frac{d\xi'}{\xi'} \int_{1/e}^1 dz \frac{\alpha_s(ze, \xi')}{2\pi} \sum_{j,k} P_{i \rightarrow jk}(z) \Phi_i(e, \xi') \left(\frac{\Phi_j(e, \xi') \Phi_k(E(z), \xi')}{\Phi_i(e, \xi')} - 1 \right),\tag{A2}$$

where the running coupling is expressed as follows [10]

$$\alpha_s(k_T^2) = \frac{\pi}{b_0 \log \frac{z^2 e^2 E_A^2 \xi}{\Lambda^2}}.\tag{A3}$$

i is either q or g and we sum over all allowed splittings. Now we take the derivative with respect to ξ and find

$$\frac{d\Phi_i(e, \xi)}{d\xi} = \frac{1}{\xi} \Phi_i(e, \xi) \int_{1/e}^1 dz \frac{\alpha_s(ze, \xi)}{2\pi} \sum_{j,k} P_{i \rightarrow jk}(z) \left(\frac{\Phi_j(e, \xi) \Phi_k(E(z), \xi)}{\Phi_i(e, \xi)} - 1 \right).\tag{A4}$$

Taking into account that $\Phi_i(e, 1) = u$ the general closed solution is

$$\Phi_i(e, \xi) = u \exp \left[\int_{\xi_R}^{\xi} \frac{d\xi'}{\xi'} \int_{1/e}^1 dz \frac{\alpha_s(ze, \xi')}{2\pi} \sum_{j,k} P_{i \rightarrow jk}(z) \left(\frac{\Phi_j(e, \xi') \Phi_k(E(z), \xi')}{\Phi_i(e, \xi')} - 1 \right) \right].\tag{A5}$$

Appendix B: Closed solution in the staircase limit

We start from Eq. (8)

$$\Phi_g(e, \xi) = u \exp \left(\int_{\xi_R}^{\xi} \frac{d\xi'}{\xi'} \int_{1/e}^1 dz \frac{\alpha_s(z, \xi')}{2\pi} P_{g \rightarrow gg}(z) \left[\Phi_g(e, \xi') + \sum_{n=1}^{\infty} \frac{(e(z-1))^n}{n!} \frac{d^n \Phi_g(e, \xi')}{de^n} - 1 \right] \right).\tag{B1}$$

Note that due its exponential form all derivatives of Φ_g can be written in the following form

$$\frac{d^n \Phi_g(e, \xi')}{de^n} = \Phi_g(e, \xi') \times \mathcal{DP}[n](e, \xi'; \Phi_g),\tag{B2}$$

where $\mathcal{DP}[n]$ is a polynomial of inner derivatives of Φ_g . We plug this in eq. (B1) and get

$$\Phi_g(e, \xi) = u \exp \left[\int_{\xi_R}^{\xi} \frac{d\xi'}{\xi'} \int_{1/e}^1 dz \frac{\alpha_s(z, \xi')}{2\pi} \times \right. \\ \left. P_{g \rightarrow gg}(z) \left(\Phi_g(e, \xi') + \underbrace{\Phi_g(e, \xi') \sum_{n=1}^{\infty} \frac{(e(z-1))^n}{n!} \mathcal{DP}[n](e, \xi'; \Phi_g) - 1}_{\mathcal{T}(z, e, \xi')} \right) \right]. \quad (\text{B3})$$

Up to this point we have not gained any new insight. We can now introduce more new symbols to rewrite eq. (B3) and bring it in the form

$$\Phi_g(e, \xi) = u \exp \left[\int_{\xi_R}^{\xi} d\xi' (\Phi_g(e, \xi') - 1) \underbrace{\frac{1}{\xi'} \int_{1/e}^1 dz \frac{\alpha_s(z, \xi')}{2\pi} P_{g \rightarrow gg}(z)}_{\gamma_g(e, \xi')} + \right. \\ \left. \int_{\xi_R}^{\xi} d\xi' \Phi_g(e, \xi') \underbrace{\frac{1}{\xi'} \int_{1/e}^1 dz \frac{\alpha_s(z, \xi')}{2\pi} P_{g \rightarrow gg}(z) \mathcal{T}(z, e, \xi')}_{r(e, \xi')} \right] \\ = u \exp \left[\int_{\xi_R}^{\xi} d\xi' (\Phi_g(e, \xi') - 1) \gamma_g(e, \xi') + \int_{\xi_R}^{\xi} d\xi' \Phi_g(e, \xi') r(e, \xi') \right]. \quad (\text{B4})$$

Taking the derivative this defines a differential equation of the form

$$\frac{d\Phi_g(e, \xi)}{d\xi} = \Phi_g(e, \xi) \times [\gamma_g(e, \xi) (\Phi_g(e, \xi) - 1) + r(e, \xi) \Phi_g(e, \xi)] \\ \Phi_g(e, \xi_r) = u. \quad (\text{B5})$$

We note that for the case $r(e, \xi) \rightarrow 0$ eq. (B5) produces exact staircase scaling [12]. Thus we connect this term to staircase breaking. We also can solve eq. (B5) for arbitrary $r(e, \xi)$. The solution is

$$\Phi_g(e, \xi) = \frac{1}{1 + \frac{(1-u)}{u\Delta_g(e, \xi)} - \underbrace{\int_{\xi_R}^{\xi} d\xi' \frac{\Delta_g(e, \xi')}{\Delta_g(e, \xi)} r(e, \xi')}_{\mathcal{R}(e, \xi)}} \quad (\text{B6})$$

where $\Delta_g(e, \xi) = \exp \left[- \int_{\xi_R}^{\xi} d\xi' \gamma(e, \xi') \right]$ is the Sudakov form factor. This is a neat result. We find that we can write the generating functional in general in a staircase like form together with an yet unspecified staircase scaling breaking term. We check the solution explicitly by differentiating eq. (B6) with respect to ξ . We note that we have

$$\frac{d\Delta_g(e, \xi)}{d\xi} = -\gamma(e, \xi) \Delta_g(e, \xi) \\ \frac{d\mathcal{R}(e, \xi)}{d\xi} = r(e, \xi) + \gamma(e, \xi) \mathcal{R}(e, \xi). \quad (\text{B7})$$

Therefore, we get

$$\begin{aligned} \frac{d\Phi_g(e, \xi)}{d\xi} &= (-1) \left(1 + \frac{1-u}{u\Delta_g(e, \xi)} - \mathcal{R}(e, \xi) \right)^{-2} \\ &\quad \times \left[\frac{1-u}{u\Delta_g(e, \xi)} \gamma(e, \xi) \Delta_g(e, \xi) - r(e, \xi) - \gamma(e, \xi) \mathcal{R}(e, \xi) \right] \\ &= \Phi_g(e, \xi) \times [\gamma_g(e, \xi) (\Phi_g(e, \xi) - 1) + r(e, \xi) \Phi_g(e, \xi)]. \end{aligned} \quad (\text{B8})$$

To evolve further we need to employ some assumptions. This means we have to drop the explicit Φ_g dependence in \mathcal{T} . For example we could expand around $e \approx 1$ and write Φ_g as an explicit series in u , dropping all higher terms in e . To find the leading u dependence we plug in all previous definitions and find that

$$\begin{aligned} \mathcal{R}(e, \xi) &= \int_{\xi_R}^{\xi} d\xi' \frac{\Delta_g(e, \xi')}{\Delta_g(e, \xi)} r(e, \xi') \\ &= \int_{\xi_R}^{\xi} d\xi' \frac{\Delta_g(e, \xi')}{\Delta_g(e, \xi)} \frac{1}{\xi'} \int_{1/e}^1 dz \frac{\alpha_s(z, \xi')}{2\pi} P_{g \rightarrow gg}(z) \mathcal{T}(z, e, \xi') \\ &= \int_{\xi_R}^{\xi} d\xi' \frac{\Delta_g(e, \xi')}{\Delta_g(e, \xi)} \frac{1}{\xi'} \int_{1/e}^1 dz \frac{\alpha_s(z, \xi')}{2\pi} P_{g \rightarrow gg}(z) \sum_{n=1}^{\infty} \frac{(e(z-1))^n}{n!} \mathcal{DP}[n](e, \xi'; \Phi_g). \end{aligned} \quad (\text{B9})$$

The task at hand is to find a Φ_g independent approximation for $\mathcal{DP}[n]$. An obvious step is to truncate the Taylor expansion at $n = 0$. We are then left with

$$\mathcal{DP}[1](e, \xi'; \Phi_g) = \frac{d\Phi_g(e, \xi')}{\Phi_g(e, \xi') de}. \quad (\text{B10})$$

We need to find the significant part in the limit $e \rightarrow 1$ and its u dependence. We have

$$\begin{aligned} \mathcal{DP}[1](e, \xi'; \Phi_g) &= \int_{\xi_R}^{\xi'} \frac{d\xi''}{\xi''} \left[\frac{\alpha_s(1/e, \xi'') P_{g \rightarrow gg}(1/e)}{2\pi e^2} \left(\underbrace{\Phi_g(1, \xi'')}_{=u} - 1 \right) + \right. \\ &\quad \left. \int_{1/e}^1 dz \frac{\alpha_s(z, \xi'') P_{g \rightarrow gg}(z) z}{2\pi} \left[\frac{d\Phi(e, \xi'')}{de} \Big|_{e=ze} \right] \right]. \end{aligned} \quad (\text{B11})$$

In principle we get a nested series of Φ_g differentiations. To argue that first term is the most important we note that formally in the $e \rightarrow 1$ limit we can write

$$\int_{1/e}^1 dz \frac{\alpha_s(z, \xi'') P_{g \rightarrow gg}(z) z}{2\pi} \left[\frac{d\Phi(e, \xi'')}{de} \Big|_{e=ze} \right] \approx (1 - 1/e) \frac{\alpha_s(1/e, \xi'') P_{g \rightarrow gg}(1/e)}{2\pi e} \left[\frac{d\Phi(e, \xi'')}{de} \Big|_{e \approx 1} \right]. \quad (\text{B12})$$

We note that the last term of eq. (B12) is exactly zero in the formal limit $e \rightarrow 1$. This is our first argument that the second term vanishes faster than the first one in the $e \rightarrow 1$ limit. We also can control its size this way and we set it to $u - 1$ which is for sure over estimated as at the end of the day we have $u \rightarrow 0$. We than see that for $e > 1$ and $e \rightarrow 1$

$$\begin{aligned} 1 &> e - 1 \\ \frac{1}{e} &> \frac{e - 1}{e} \\ \frac{1}{e} &> (1 - 1/e). \end{aligned} \quad (\text{B13})$$

Therefore, the first term is greater than the second one. However, this is only valid in the $e \rightarrow 1$ limit. From this estimate we can not deduce the uncertainty we introduce by dropping the second term nor can we conclude on

the range of e where this estimate is valid, because we do not know the correct e dependence. Nevertheless, we compute formally the u dependence of the staircase breaking term in the staircase limit $e \rightarrow 1$. Thus we write

$$\begin{aligned}
\mathcal{DP}[1](e, \xi'; \Phi_g) &\approx (u-1) \int_{\xi_R}^{\xi'} \frac{d\xi''}{\xi''} \frac{\alpha_s(1/e, \xi'') P_{g \rightarrow gg}(1/e)}{2\pi e^2} (u-1) \\
&\equiv \rho(e, \xi') \\
\mathcal{R}(e, \xi) &\approx (u-1) \int_{\xi_R}^{\xi} d\xi' \frac{\Delta_g(e, \xi')}{\Delta_g(e, \xi)} \frac{1}{\xi'} \int_{1/e}^1 dz \frac{\alpha_s(z, \xi')}{2\pi} P_{g \rightarrow gg}(z) (z-1) e(u-1) \rho(e, \xi') \\
&\equiv \chi(e, \xi).
\end{aligned} \tag{B14}$$

-
- [1] T. Sjostrand, S. Mrenna, and P. Z. Skands, “PYTHIA 6.4 Physics and Manual,” *JHEP*, vol. 0605, p. 026, 2006.
 - [2] M. Bahr, S. Gieseke, M. Gigg, D. Grellscheid, K. Hamilton, *et al.*, “Herwig++ Physics and Manual,” *Eur.Phys.J.*, vol. C58, pp. 639–707, 2008.
 - [3] T. Gleisberg, S. Höche, F. Krauss, A. Schälicke, S. Schumann, and J. Winter, “SHERPA 1.α, a proof-of-concept version,” *JHEP*, vol. 02, p. 056, 2004.
 - [4] S. Catani, B. R. Webber, and G. Marchesini, “QCD coherent branching and semiinclusive processes at large x ,” *Nucl. Phys.*, vol. B349, pp. 635–654, 1991.
 - [5] S. D. Ellis and D. E. Soper, “Successive combination jet algorithm for hadron collisions,” *Phys. Rev.*, vol. D48, pp. 3160–3166, 1993.
 - [6] S. Catani, Y. L. Dokshitzer, M. Olsson, G. Turnock, and B. R. Webber, “New clustering algorithm for multijet cross sections in e^+e^- annihilation,” *Phys. Lett.*, vol. B269, pp. 432–438, 1991.
 - [7] M. Cacciari, G. P. Salam, and G. Soyez, “The Anti- $k(t)$ jet clustering algorithm,” *JHEP*, vol. 04, p. 063, 2008.
 - [8] Y. L. Dokshitzer, G. Leder, S. Moretti, and B. Webber, “Better jet clustering algorithms,” *JHEP*, vol. 9708, p. 001, 1997.
 - [9] M. Wobisch and T. Wengler, “Hadronization corrections to jet cross-sections in deep inelastic scattering,” 1998.
 - [10] E. Gerwick, S. Schumann, B. Gripaios, and B. Webber, “QCD Jet Rates with the Inclusive Generalized kt Algorithms,” *JHEP*, vol. 1304, p. 089, 2013.
 - [11] S. Catani, Y. L. Dokshitzer, F. Fiorani, and B. Webber, “Average number of jets in e^+e^- annihilation,” *Nucl.Phys.*, vol. B377, pp. 445–460, 1992.
 - [12] E. Gerwick, T. Plehn, S. Schumann, and P. Schichtel, “Scaling Patterns for QCD Jets,” *JHEP*, vol. 1210, p. 162, 2012.
 - [13] E. Gerwick, T. Plehn, and S. Schumann, “Understanding Jet Scaling and Jet Vetos in Higgs Searches,” *Phys.Rev.Lett.*, vol. 108, p. 032003, 2012.
 - [14] M. Dasgupta, A. Fregoso, S. Marzani, and G. P. Salam, “Towards an understanding of jet substructure,” *JHEP*, vol. 1309, p. 029, 2013.
 - [15] M. Dasgupta, A. Fregoso, S. Marzani, and A. Powling, “Jet substructure with analytical methods,” *Eur.Phys.J.*, vol. C73, no. 11, p. 2623, 2013.
 - [16] M. Dasgupta, F. Dreyer, G. P. Salam, and G. Soyez, “Small-radius jets to all orders in QCD,” 2014.
 - [17] S. Alioli and J. R. Walsh, “Jet Veto Clustering Logarithms Beyond Leading Order,” *JHEP*, vol. 1403, p. 119, 2014.
 - [18] F. J. Tackmann, J. R. Walsh, and S. Zuberi, “Resummation Properties of Jet Vetoes at the LHC,” *Phys.Rev.*, vol. D86, p. 053011, 2012.
 - [19] S. D. Ellis, C. K. Vermilion, J. R. Walsh, A. Hornig, and C. Lee, “Jet Shapes and Jet Algorithms in SCET,” *JHEP*, vol. 1011, p. 101, 2010.
 - [20] G. Aad *et al.*, “Measurement of the production cross section of jets in association with a Z boson in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector,” 2013.
 - [21] C. Englert, T. Plehn, P. Schichtel, and S. Schumann, “Establishing Jet Scaling Patterns with a Photon,” *JHEP*, vol. 1202, p. 030, 2012.
 - [22] Z. Bern *et al.*, “Universality in W+Multijet Production,” *PoS*, vol. LL2014, pp. 283–288, 2014.
 - [23] Z. Bern *et al.*, “Universality in W+Multijet Production,” pp. 283–288, 2014.
 - [24] C. Englert, T. Plehn, P. Schichtel, and S. Schumann, “Jets plus Missing Energy with an Autofocus,” *Phys. Rev.*, vol. D83, p. 095009, 2011.
 - [25] S. El Hedri, A. Hook, M. Jankowiak, and J. G. Wacker, “Learning How to Count: A High Multiplicity Search for the LHC,” *JHEP*, vol. 1308, p. 136, 2013.
 - [26] E. Gerwick, “Recursive prescription for logarithmic jet rate coefficients,” *Phys.Rev.*, vol. D88, no. 9, p. 094009, 2013.
 - [27] M. Cacciari, G. P. Salam, and G. Soyez, “The Catchment Area of Jets,” *JHEP*, vol. 0804, p. 005, 2008.
 - [28] M. Cacciari, G. P. Salam, and G. Soyez, “FastJet user manual,” *Eur.Phys.J.*, vol. C72, p. 1896, 2012.

- [29] T. Gleisberg, S. Höche, F. Krauss, M. Schönherr, S. Schumann, F. Siegert, and J. Winter, “Event generation with Sherpa 1.1,” *JHEP*, vol. 02, p. 007, 2009.
- [30] S. Schumann and F. Krauss, “A parton shower algorithm based on Catani-Seymour dipole factorisation,” *JHEP*, vol. 03, p. 038, 2008.
- [31] P. Pfeifenschneider *et al.*, “QCD analyses and determinations of α_s in e^+e^- annihilation at energies between 35-GeV and 189-GeV,” *Eur. Phys. J.*, vol. C17, pp. 19–51, 2000.
- [32] G. Aad *et al.*, “Measurement of kT splitting scales in W- $\ell\nu$ events at $\sqrt{s}=7$ TeV with the ATLAS detector,” *Eur.Phys.J.*, vol. C73, p. 2432, 2013.
- [33] S. Catani, Y. L. Dokshitzer, M. H. Seymour, and B. R. Webber, “Longitudinally-invariant k_\perp -clustering algorithms for hadron-hadron collisions,” *Nucl. Phys.*, vol. B406, pp. 187–224, 1993.
- [34] R. K. Ellis, W. J. Stirling, and B. R. Webber, *QCD and collider physics*, vol. 8. Cambridge Monogr. Part. Phys. Nucl. Phys. Cosmol., 1 ed., 1996.